Statistics of sums of correlated variables described by a matrix product ansatz

Florian Angeletti, Eric Bertin, and Patrice Abry
Université de Lyon, Laboratoire de Physique, ENS Lyon, CNRS, 46 Allée d'Italie, F-69007 Lyon, France
(Dated: April 22, 2013)

We determine the asymptotic distribution of the sum of correlated variables described by a matrix product ansatz with finite matrices, considering variables with finite variances. In cases when the correlation length is finite, the law of large numbers is obeyed, and the rescaled sum converges to a Gaussian distribution. In constrast, when correlation extends over system size, we observe either a breaking of the law of large numbers, with the onset of giant fluctuations, or a generalization of the central limit theorem with a family of nonstandard limit distributions. The corresponding distributions are found as mixtures of delta functions for the generalized law of large numbers, and as mixtures of Gaussian distributions for the generalized central limit theorem. Connections with statistical physics models are emphasized.

PACS numbers: 05.40.-a, 02.50.Ey, 47.27.eb

The law of large numbers and the central limit theorem are cornerstones of equilibrium statistical physics. Indeed, the very existence of determistic values of macroscopic observables in large systems relies on the law of large numbers, while the Gaussian shape of tiny fluctuations around the mean value are described by the central limit theorem. The latter also bears strong connections with random walks [1]. Basic forms of these theorems are known for independent and identically distributed (i.i.d.) random variables [1–3]. Some generalizations with less restrictive assumptions are also known [3, 4], and it is often assumed that these theorems remain valid as long as distributions of individual variables are not too broad and do not differ too much one from the other, and as long as correlations are weak enough. These assumptions are however not always valid in nonequilibrium systems, where the relevant observables may have an infinite mean (e.g., in aging [5, 6] or laser cooling [7] phenomena), or may have long-range correlations (e.g., in boundary driven [8, 9] or active [10, 11] systems), leading to the breakdown of the law of large numbers. Similarly, the standard central limit theorem breaks down in a number of cases. For broadly distributed variables with infinite variance, the generalized central limit theorem yields Lévy-stable laws [2], with many applications often related to anomalous diffusion [5]. Non-Gaussian distributions have also been found for instance in the context of $1/f^\alpha$-noise [12] and related problems [13, 14], where summed variables have very different statistics. For strongly correlated variables, generalizations of the central limit theorem have been derived for Gaussian processes [15, 16]. However, in a statistical physics context, another class of strongly correlated variables, defined through a matrix product ansatz, has emerged from the exact solution of nonequilibrium models like the Asymmetric Simple Exclusion Process (ASEP) [17, 18]. Although infinite matrices may be needed, notably in the context of the ASEP model [17, 18], a significant number of models can be solved using finite matrices, including reaction-diffusion lattice models [18–20], and the ASEP model for specific parameter values [9, 21]. In spite of the increasing importance of this class of random variables, the corresponding generalizations of the law of large numbers and of the central limit theorem are presently not known. In this Letter, we aim at providing such generalizations for variables described by a matrix product ansatz with finite matrices. Our study encompasses both discrete random variables as in the ASEP model [17] and continuous ones as in signal processing [22, 23] or in mass transport models [24]. For the sake of clarity, we restrict our presentation to specific, yet representative cases, deferring a full-length account of our results to a forthcoming publication [25].

Matrix product representation.– Following [22, 23], we consider a set of random variables $(x_1, \ldots, x_N)$ whose joint probability distribution can be described by a matrix product ansatz, namely

$$P(x_1, \ldots, x_N) = \frac{1}{\mathcal{L}(\mathcal{E}^N)} \mathcal{L}\big(\mathcal{R}(x_1)\mathcal{R}(x_2)\ldots\mathcal{R}(x_N)\big) \quad (1)$$

where $\mathcal{R}(x)$ is a $D \times D$ matrix function ($D \geq 2$) with real nonnegative entries, $\mathcal{E} = \int_{-\infty}^{\infty} \mathcal{R}(x)\, dx$, and $\mathcal{L}$ is a linear form defined as

$$\mathcal{L}\,(M) = \mathsf{tr}\left(\mathcal{A}^T M\right), \quad (2)$$

with $\mathcal{A}$ a given $D \times D$, nonzero matrix with real nonnegative entries. We further assume that for all $N \geq 1$, $\mathcal{L}(\mathcal{E}^N) \neq 0$. Standard forms used in statistical physics for the linear form $\mathcal{L}$ [18] are recovered either by taking $\mathcal{A}$ as the identity matrix, or by choosing $\mathcal{A}_{ij} = V_i W_j$ so that $\mathcal{L}(M) = \langle V|M|W\rangle$. Eq. (1) is a natural generalization to correlated variables of the i.i.d. case, replacing the product of real functions by a product of matrix functions. It is useful to introduce the matrix of distributions $\mathcal{P}(x)$ through the relation

$$\mathcal{R}_{ij}(x) = \mathcal{E}_{ij}\mathcal{P}_{ij}(x), \quad (3)$$

so that $\mathcal{P}_{ij}(x)$ can be interpreted as a probability distribution, normalized to 1. We consider probabilities $\mathcal{P}_{ij}(x)$

with finite mean value m_{ij} and finite variance σ_{ij}^2. Note that $\mathcal{P}_{ij}(x)$ is uniquely defined only when $\mathcal{E}_{ij} \neq 0$.

Hidden Markov Chain representation.– As shown in [22, 23], the joint probability (1) can be reinterpreted using the concept of Hidden Markov Chain [26]. Expanding the matrix product, one can rewrite Eq. (1) as

$$P(x_1, \ldots, x_N) = \sum_{\Gamma} \kappa_\Gamma \, P_\Gamma(x_1, \ldots, x_N) \tag{4}$$

with $\Gamma = (\Gamma_1, \ldots, \Gamma_{N+1})$ a chain of indices, and

$$\kappa_\Gamma = \frac{\mathcal{A}_{\Gamma_1 \Gamma_{N+1}}}{\mathcal{L}\left(\mathcal{E}^N\right)} \mathcal{E}_{\Gamma_1 \Gamma_2} \ldots \mathcal{E}_{\Gamma_N \Gamma_{N+1}}, \tag{5}$$

$$P_\Gamma(x_1, \ldots, x_N) = \prod_{k=1}^{N} \mathcal{P}_{\Gamma_k \Gamma_{k+1}}(x_k). \tag{6}$$

Eq. (4) can be interpreted as a mixture of distributions $P_\Gamma(x_1, \ldots, x_N)$ indexed by the chain of indices Γ, each with probability weight κ_Γ (note that $\sum_\Gamma \kappa_\Gamma = 1$). For a given Γ, the distribution $P_\Gamma(x_1, \ldots, x_N)$ describes independent, but nonidentically distributed random variables. Correlations, when present, thus emerge from the mixture of the different distributions P_Γ.

Interestingly, the probability weight κ_Γ can be interpreted as the probability of a Markov chain [22, 23]. More precisely, given a starting point Γ_1 and a final point Γ_{N+1}, κ_Γ can be recast as the probability of an inhomogeneous Markov chain with transition probabilities

$$p(\Gamma_{k+1} = j | \Gamma_k = i, \, \Gamma_{N+1} = f) = \mathcal{E}_{ij} \frac{(\mathcal{E}^{N-k})_{jf}}{(\mathcal{E}^{N-k+1})_{if}}. \tag{7}$$

Note that this Markov chain is of a nonstandard type, due to the dependence on the final state Γ_{N+1}. Both the initial and final states thus need to be selected in advance, and their joint distribution is given by:

$$\tilde{p}(\Gamma_1 = i, \Gamma_{N+1} = f) = \mathcal{A}_{if} \frac{(\mathcal{E}^N)_{if}}{\mathcal{L}(\mathcal{E}^N)}. \tag{8}$$

Hence the procedure to simulate the correlated random variables described by Eq. (4) is the following [23]: (i) Γ_1 and Γ_{N+1} are chosen at random according to distribution (8); (ii) the random chain Γ is obtained from transition rates (7); (iii) the random variables x_k ($k = 1, \ldots N$) are drawn randomly from the distributions $\mathcal{P}_{\Gamma_k \Gamma_{k+1}}(x_k)$.

As seen in Eq. (7), the matrix $\mathcal{E}$ plays a key role in statistical properties of the Markov chain Γ, which in turn determine correlations between the x_k's. It can be shown in particular (see [23] for examples) that $\mathcal{E}$ determines the short-range or long-range nature of the correlations. If the eigenvalue λ_1 of $\mathcal{E}$ with the largest modulus has a multiplicity of 1, the correlation $C_{kl} = \langle x_k x_l \rangle - \langle x_k \rangle \langle x_l \rangle$ is quasi-stationary, short-ranged and decays exponentially to zero at large $|k - l|$. A prominent class of matrices obeying this property is the class of aperiodic irreducible matrices (see below), due to Perron-Frobenius theorem [27]. In contrast, if the multiplicity of λ_1 is strictly larger than 1, the correlation C_{kl} (or more precisely, the full dependence structure) is generically long-ranged. Two situations may arise in this case, according to the Jordan decomposition of the matrix $\mathcal{E}$ [28]. If λ_1 is associated to a diagonal block, the correlation C_{kl} decays to a constant (and generically nonzero) value for large $|k - l|$. If λ_1 is associated to a nondiagonal block, the correlation C_{kl} is nonstationary, and becomes a function of k/N and l/N, so that correlations also extend over system size [9, 23]. We come back below to these different classes of matrices.

Statistics of the sum.– We now focus on the study of the statistical properties of the sum $S = \sum_{k=1}^{N} x_k$, in the limit $N \to \infty$. We are specifically interested in the validity of the law of large numbers and of the central limit theorem. Introducing the variable $s = S/N$, the law of large numbers breaks down if the limit distribution $\Psi(s)$ for $N \to \infty$ does not reduce to a delta function. When the law of large numbers holds, namely $\Psi(s) = \delta(s - m)$, one can investigate the fluctuations of S around its mean value Nm on a scale $\sim \sqrt{N}$, through the rescaled variable

$$z = \frac{S - Nm}{\sqrt{N}}. \tag{9}$$

When correlations are weak enough, the central limit theorem should hold, and the distribution $\Phi_N(z)$ should converge to a Gaussian when $N \to \infty$.

The Hidden Markov Chain formalism proves very useful in order to study the statistical properties of the sum, as it allows to separate two different sources of randomness: the random choice of the chain Γ, and the random choice of $(x_1, \ldots, x_N)$ from the distribution P_Γ, for a fixed Γ. Hence the distribution of S can be determined by first computing the distribution of $S_\Gamma = \sum_{k=1}^{N} x_k$, where the x_k's are drawn from P_Γ, and then averaging the distribution of S_Γ over Γ. For a given chain Γ, we introduce the fraction ν_{ij} of transitions from i to j in Γ,

$$\nu_{ij} = \frac{1}{N} \, \mathrm{card}\{k, \Gamma_k = i \text{ and } \Gamma_{k+1} = j\}. \tag{10}$$

For a given Γ, the sum S_Γ can be rewritten as

$$S_\Gamma = \sum_{i,j=1}^{D} \sum_{k=1}^{N\nu_{ij}} X_k^{(ij)}, \tag{11}$$

where the variables $X_k^{(ij)}$, $k = 1, \ldots, N\nu_{ij}$, are i.i.d. random variables drawn from the distribution $\mathcal{P}_{ij}$. Let us first investigate the validity of the law of large numbers for the sum S. For a fixed Γ, one can write

$$\frac{S_\Gamma}{N} = \sum_{i,j=1}^{D} \nu_{ij} \left(\frac{1}{N\nu_{ij}} \sum_{k=1}^{N\nu_{ij}} X_k^{(ij)} \right). \tag{12}$$

Given that, for fixed (i,j), the variables $X_k^{(ij)}$ are i.i.d., the law of large numbers can be applied to the term between brackets, which thus converges (almost surely) to the mean value m_{ij}. Hence, for a fixed $\nu = (\nu_{ij})$, the conditional distribution of s converges when $N \to \infty$ to

$$\Psi(s|\nu) = \delta\left(s - \sum_{i,j=1}^{D} \nu_{ij} m_{ij}\right). \tag{13}$$

As the only dependence over Γ is through ν, the average over Γ can be replaced by an average over ν. The limit distribution $\Psi(s)$ is thus obtained by averaging $\Psi(s|\nu)$ over the asymptotic ($N \to \infty$) distribution of ν, denoted as $Q(\nu)$:

$$\Psi(s) = \int \prod_{i,j=1}^{D} d\nu_{ij}\, Q(\nu)\, \delta\left(s - \sum_{i,j=1}^{D} \nu_{ij} m_{ij}\right). \tag{14}$$

The law of large numbers holds either when all m_{ij}'s are equal (or at least those associated to nonzero ν_{ij}), or when the empirical frequencies ν_{ij} converge to non-random values $\overline{\nu}_{ij}$ in the limit $N \to \infty$, in which case the rescaled sum s converges to the deterministic limit $m = \sum_{i,j} \overline{\nu}_{ij} m_{ij}$. When the law of large numbers is satisfied, the validity of the central limit theorem can then be investigated. To this aim, we use the variable z defined in Eq. (9), and follow a similar path as above. For a given chain Γ, and thus a given ν, one has

$$\frac{S_\Gamma - Nm}{\sqrt{N}} = \sum_{i,j=1}^{D} \sqrt{\nu_{ij}} \left(\frac{\sum_{k=1}^{N\nu_{ij}} X_k^{(ij)} - N\nu_{ij} m_{ij}}{\sqrt{N\nu_{ij}}}\right). \tag{15}$$

As for fixed (i,j), the variables $X_k^{(ij)}$ are i.i.d. with finite variance, the central limit theorem applies to the sum between brackets, and the distribution of this sum converges, for $N \to \infty$, to a centered Gaussian distribution of variance σ_{ij}. Hence the conditional distribution $\Phi(z|\nu)$ is a centered Gaussian distribution of variance $\sum_{i,j} \nu_{ij} \sigma_{ij}^2$. Averaging over ν yields the distribution

$$\Phi(z) = \int \frac{\prod_{i,j} d\nu_{ij}\, Q(\nu)}{\sqrt{2\pi \sum_{i,j} \nu_{ij} \sigma_{ij}^2}}\, e^{-z^2/(2\sum_{i,j} \nu_{ij}\sigma_{ij}^2)}. \tag{16}$$

The central limit theorem then holds either when all σ_{ij}'s (associated to nonzero ν_{ij}) are equal, or when the ν_{ij}'s take deterministic values. When the central limit theorem does not hold, the distribution $\Phi(z)$ is a mixture of Gaussian distributions.

As seen above, the distribution $Q(\nu)$ is thus the key ingredient to characterize the limit distributions $\Psi(s)$ and (when the law of large numbers holds) $\Phi(z)$. In the following, we determine $Q(\nu)$ in representative cases, yielding typical examples of limit distributions.

Irreducible matrix $\mathcal{E}$.– Let us first consider the case where the matrix $\mathcal{E}$ is irreducible. To check this property, given that $\mathcal{E}$ is nonnegative, one can consider the associated directed graph: introducing D nodes, a directed link is defined from node i to node j if $\mathcal{E}_{ij} > 0$. The matrix $\mathcal{E}$ is said to be irreducible if the graph is strongly connected (which means that for any (i,j), there exists a directed path from i to j). If $\mathcal{E}$ is irreducible and aperiodic [31], the Perron-Frobenius theorem [27] implies that $\mathcal{E}$ admits a dominant positive eigenvalue λ_1, namely all other eigenvalues λ_i are such that $|\lambda_i| < \lambda_1$. The eigenvector $\mu_i^{(1)}$ associated to λ_1 has positive components, and can be normalized such that $\sum_i \mu_i^{(1)} = 1$. Apart from boundary effects appearing when $N - k \ll N$, the transition probability of the Hidden Markov Chain becomes homogeneous, that is, independent of k and of the final state f,

$$p(\Gamma_{k+1} = j | \Gamma_k = i) \approx \mathcal{E}_{ij} \frac{\mu_j^{(1)}}{\lambda_1 \mu_i^{(1)}}. \tag{17}$$

From standard results of homogeneous Markov Chain theory [27], the chain converges to a unique stationary state described by a probability $q_i = \mu_i^{(1)} \rho_i^{(1)}$ to occupy state i, where $\rho^{(1)}$ is the λ_1-eigenvector of $\mathcal{E}^T$. The convergence towards the stationary state implies that, for $N \to \infty$, ν_{ij} converges (almost surely) to the average value $\overline{\nu}_{ij} = q_i \mathcal{E}_{ij} \mu_j^{(1)} / \lambda_1 \mu_i^{(1)}$. The empirical frequencies ν_{ij} thus do not fluctuate and both the law of large numbers and the central limit theorem hold. This can be seen formally by plugging the distribution $Q(\nu) = \prod_{i,j} \delta(\nu_{ij} - \overline{\nu}_{ij})$ into Eqs. (14) and (16). Interestingly, this result is independent of the form of the matrix $\mathcal{A}$. Note also that the above result is consistent with the fact that correlations are short-ranged when the matrix $\mathcal{E}$ has a single dominant eigenvalue.

Diagonalizable reducible matrix $\mathcal{E}$.– Nonstandard distributions are expected to emerge when correlation extends over system size. As mentioned above, this is true when the multiplicity of the largest eigenvalue of the matrix $\mathcal{E}$ is strictly larger than 1, which implies that the matrix is reducible. A simple example of such a situation is when the multiplicity is equal to D. Assuming $\mathcal{E}$ to be diagonalizable, it is then clear that $\mathcal{E}$ is proportional to the identity matrix I. The proportionality factor can be scaled out by a redefinition of $\mathcal{R}(x)$ and thus of $\mathcal{E}$ [see Eq. (1)], so that we set $\mathcal{E} = I$. From Eq. (5), one sees that the only possible chains are the constant chains $\Gamma^{(i)} = (i, i, \ldots, i)$, for $i = 1, \ldots D$. For a given chain $\Gamma^{(i)}$, one thus has $\nu_{ii} = 1$ and $\nu_{kl} = 0$ for $(k,l) \neq (i,i)$. In addition, Eq. (8) implies that the chain $\Gamma^{(i)}$ appears with probability $q_i = \mathcal{A}_{ii} / \sum_j \mathcal{A}_{jj}$. As a result, one finds

$$Q(\nu) = \frac{1}{D} \sum_{i=1}^{D} q_i\, \delta(\nu_{ii} - 1) \left[\prod_{(k,l)\neq(i,i)} \delta(\nu_{kl})\right], \tag{18}$$

which, from Eq. (14), leads to a generalization of the law

of large numbers,

$$\Psi(s) = \frac{1}{D}\sum_{i=1}^{D} q_i\,\delta\left(s - m_{ii}\right). \tag{19}$$

If all the m_{ii}'s are equal, the standard law of large numbers holds, and using Eq. (16), the central limit theorem is generalized into

$$\Phi(z) = \frac{1}{D}\sum_{i=1}^{D} \frac{q_i}{\sqrt{2\pi\sigma_{ii}^2}}\, e^{-z^2/2\sigma_{ii}^2}. \tag{20}$$

Nondiagonalizable reducible matrix $\mathcal{E}$.– Remaining in the framework of reducible matrices $\mathcal{E}$ with a multiple dominant eigenvalue (so that correlation is generically long-ranged), we turn to the nondiagonalizable case. A simple and representative example of this case consists in considering a Jordan block,

$$\mathcal{E} = \lambda I + U, \quad U = \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix}. \tag{21}$$

Through both a change of basis and a rescaling of $\mathcal{E}$, one can set $\lambda = 1$. From Eq. (5), chains Γ having nonzero probability contain only transitions from i to $i+1$. If $\mathcal{A}_{1n} \neq 0$, it can be shown that with probability 1, Γ starts from $\Gamma_1 = 1$ and end at $\Gamma_{N+1} = D$, in the limit $N \to \infty$. Moreover, Eq. (5) implies that all such chains are equiprobable. Given that for $i \neq j$, these chains have at most one transition from i to j, the corresponding empirical frequencies ν_{ij} converge to zero for $N \to \infty$. Only the frequencies ν_{ii} are nonzero, and the equiprobability of chains implies that they are uniformly distributed, under the constraint $\sum_i \nu_{ii} = 1$. Hence, at odds with previous cases, empirical frequencies have a continuous distribution in the limit $N \to \infty$,

$$Q(\nu) = \frac{1}{(D-1)!}\,\delta\left(\sum_{i=1}^{D}\nu_{ii} - 1\right)\prod_{k\neq l}\delta\left(\nu_{kl}\right). \tag{22}$$

From Eq. (14), the generalized law of large numbers reads

$$\Psi(s) = \int \frac{\prod_{i=1}^{D} d\nu_{ii}}{(D-1)!}\,\delta\left(\sum_{i=1}^{D}\nu_{ii} - 1\right)\delta\left(s - \sum_{i=1}^{D}\nu_{ii}m_{ii}\right). \tag{23}$$

The distribution $\Psi(s)$ is piecewise polynomial, with support $[\min\{m_{ii}\}, \max\{m_{ii}\}]$ (see below for an explicit example). If all the m_{ii}'s are equal, the standard law of large numbers holds, and a generalized form of the central limit theorem is obtained from Eq. (16) as

$$\Phi(z) = \int \frac{\prod_i d\nu_{ii}}{(D-1)!}\,\delta\left(\sum_i \nu_{ii} - 1\right)\frac{e^{-z^2/(2\sum_i \nu_{ii}\sigma_{ii}^2)}}{\sqrt{2\pi\sum_i \nu_{ii}\sigma_{ii}^2}}. \tag{24}$$

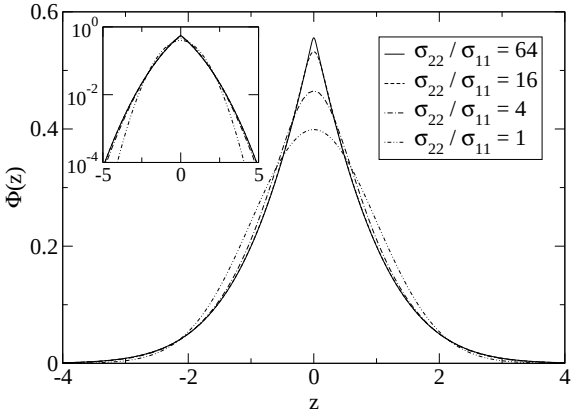

FIG. 1: *Limit distribution $\Phi(z)$ generalizing the Gaussian distribution in the case of two-dimensional matrices –see Eq.* (26)*, for different ratios σ_{11}/σ_{22}. Parameters are chosen such that all distributions have variance* 1*. The Gaussian distribution is recovered for $\sigma_{11} = \sigma_{22}$. Inset: same data on a semi-logarithmic scale.*

More generally, for an arbitrary matrix $\mathcal{E}$, an analysis can be performed in terms of decomposition of $\mathcal{E}$ into irreducible blocks. Ergodicity is then achieved within a given block, and irreversible transitions (similar to the transitions in the simple nondiagonalizable case above) take place. The final result for the distribution $\Psi(s)$ (or when applicable, for $\Phi(z)$) is generically a complicated mixture, both continuous and discrete, of standard laws. Details will be given in a forthcoming publication [25].

Two-dimensional matrix $\mathcal{E}$.– To give a more explicit example, we now consider the case of a two-dimensional matrix $\mathcal{E}$. Nonstandard limit distributions emerge when correlation extends over system size, which for $D = 2$ happens only if $\mathcal{E}$ takes the form

$$\mathcal{E} = \begin{pmatrix} \lambda & \gamma \\ 0 & \lambda \end{pmatrix}, \qquad \lambda > 0, \quad \gamma \geq 0, \tag{25}$$

or its transpose. If $\gamma = 0$, $\mathcal{E} = \lambda I$, and the limit distributions are given by Eqs. (19) and (20) with $D = 2$. If $\gamma > 0$, $\mathcal{E}$ is nondiagonalizable and the limit distributions are given by Eqs. (23) and (24), which take a relatively simple form for $D = 2$. Indeed, $\Psi(s)$ is simply a uniform distribution on the interval $[\min\{m_{ii}\}, \max\{m_{ii}\}]$, and if $m_{11} = m_{22}$,

$$\Phi(z) = \sqrt{\frac{2}{\pi}}\int_{\sigma_{11}}^{\sigma_{22}} \frac{d\sigma}{\sigma_{22}^2 - \sigma_{11}^2}\, e^{-z^2/2\sigma^2}. \tag{26}$$

This distribution is illustrated in Fig. 1 for different values of σ_{22}/σ_{11}, keeping the variance of $\Phi(z)$ fixed to 1. Increasing σ_{22}/σ_{11} makes the central peak of the distribution sharper, while the tails remain essentially Gaussian.

Discussion.– It is interesting to try to relate the above results to the behavior of some statistical physics models. For a two-dimensional nondiagonalizable matrix,

we have seen above that the chain Γ takes the form $\Gamma = (1, \ldots, 1, 2, \ldots, 2)$, with a single transition from state 1 to state 2. This property can be illustrated within the two-dimensional representation of the ASEP model [9, 21], in which $x_k = 0$ or 1 characterizes the occupancy of site k. Although Γ cannot be directly observed, it can be interpreted as the origin of the front separating high and low density regions on the coexistence line of the ASEP model [29]. The fact that the position of the front is uniformly distributed over the system [29, 30] precisely reflects the random transition in the chain Γ, and results in a uniform distribution $\Psi(s)$ on $[\rho_{\min}, \rho_{\max}]$. Interestingly, this two-dimensional picture also seems qualitatively valid when matrix representations of higher dimensions (either finite or infinite) are needed, since in all cases, a single front is present. For finite dimensional representations, this property is understood from the fact that the Jordan block associated to the dominant eigenvalue is two-dimensional [21], while the situation is less clear for infinite matrices. In any case, it would be interesting to try to extend the present approach to some relevant classes of infinite dimensional matrices.

[1] W. Feller, *An Introduction to Probability Theory and its Applications, Vol. I* (Wiley, New York, 1966).

[2] B. V. Gnedenko and A. N. Kolmogorov, *Limit Distributions for Sums of Independent Random Variables* (Addison-Wesley, Cambridge, USA, 1954).

[3] W. Feller, *An Introduction to Probability Theory and its Applications, Vol. II* (Wiley, New York, 1966).

[4] V. V. Petrov, *Limit Theorems of Probability Theory* (Oxford University Press, Oxford, 1995).

[5] J.-P. Bouchaud and A. Georges, Phys. Rep. **195**, 127 (1990).

[6] J.-P. Bouchaud, J. Phys. I (France) **2**, 1705 (1992).

[7] F. Bardou, J.-P. Bouchaud, A. Aspect, and C. Cohen-Tannoudji, *Lévy Statistics and Laser Cooling* (Cambridge University Press, Cambridge, 2002).

[8] B. Derrida and M. R. Evans, J. Phys. I (France) **3**, 311 (1993).

[9] F. H. L. Essler and V. Rittenberg, J. Phys. A: Math. Gen. **29**, 3375 (1996).

[10] S. Ramaswamy, R. A. Simha, and J. Toner, Europhys. Lett. **62**, 196 (2003).

[11] J. Toner, Y. Tu, and S. Ramaswamy, Ann. Phys. (Amsterdam) **318**, 170 (2005).

[12] T. Antal, M. Droz, G. Györgyi, and Z. Rácz, Phys. Rev. E **65**, 046140 (2002).

[13] S. T. Bramwell, J.-Y. Fortin, P. C. W. Holdsworth, S. Peysson, J.-F. Pinton, B. Portelli, and M. Sellitto, Phys. Rev. E **63**, 041106 (2001).

[14] M. Clusel and E. Bertin, Int. J. Mod. Phys. B **22**, 3311 (2008).

[15] M. S. Taqqu, Z. Wahrsch. Verw. Gebiete **50**, 53 (1979).

[16] M. Rosenblatt, Z. Wahrsch. Verw. Gebiete **55**, 123 (1981).

[17] B. Derrida, M. R. Evans, V. Hakim, and V. Pasquier, J Phys. A **26**, 1493 (1993).

[18] R. A. Blythe and M. R. Evans, Journal of Physics A: Mathematical and Theoretical **40**, R333 (2007).

[19] H. Hinrichsen, S. Sandow, and I. Peschel, J. Phys. A: Math. Gen. **29**, 2643 (1996).

[20] U. Basu and P. K. Mohanty, Phys. Rev. E **79**, 041143 (2009).

[21] K. Mallick and S. Sandow, J. Phys. A: Math. Gen. **30**, 4513 (1997).

[22] F. Angeletti, E. Bertin, and P. Abry, in *Proceedings of the IEEE ICASSP Conference* (2012), p. 3897.

[23] F. Angeletti, E. Bertin, and P. Abry (2012), arxiv:1203.4500.

[24] E. Bertin, in preparation (2013).

[25] F. Angeletti, E. Bertin, and P. Abry, in preparation (2013).

[26] O. Cappe, E. Moulines, and T. Ryden, *Inference in Hidden Markov Models*, Springer Series in Statistics (Springer, New York, 2005).

[27] E. Seneta, *Non-negative Matrices and Markov Chains*, Springer Series in Statistics (Springer, New-York, 2006).

[28] R. Horn and C. Johnson, *Matrix Analysis* (Cambridge University Press, 1990).

[29] G. Schütz and E. Domany, J. Stat. Phys. **72**, 277 (1993).

[30] M. Depken and R. Stinchcombe, Phys. Rev. Lett. **93**, 040602 (2004).

[31] A nonnegative irreducible matrix is aperiodic if for all n above some threshold, $(\mathcal{E}^n)_{ij} > 0$ for all (i, j). A sufficient condition (for a nonnegative irreducible matrix) is that at least one diagonal term is nonzero.